MILLIMAN WHITE PAPER

Impact of credit data for the valuation of insurance liabilities

Pierre-Edouard Arrouy
Emmanuel Avril
David Baranes[1]
Adrien Cortes
Rashi Garg[2]
Léo Tondolo

This paper investigates how the choice of financial data can impact the calibration and the simulation of credit spread and default economic scenarios within an economic scenario generator (ESG) as well as the insurance liability valuation metrics.

The use of market-consistent scenarios is crucial for the assessment of insurance liabilities within the regulatory frameworks of Solvency II, International Financial Reporting Standard (IFRS) 17, Long Duration Targeted Improvements (LDTI) and certain risk-based capital (RBC) regimes in Asia.

To this extent, the integration of the major financial risk exposure in the valuation procedure of the balance sheet is required by the regulators, as illustrated below for Solvency II:[3]

"Insurance and reinsurance undertakings should be able to demonstrate that the choice of financial instruments used in the calibration process is relevant given the characteristics of [their] obligations."

For this reason, an increasing number of insurers are now embedding credit risk in their economic scenario generator (ESG). The calibration of the risk-neutral credit models often targets benchmark market data spreads extracted from external data providers. The European Insurance and Occupational Pensions Authority (EIOPA) underlines the necessity of monitoring the accuracy, the appropriateness and the completeness of such data:

"To carry out the assessment of the level of accuracy, appropriateness and completeness of external data, insurance and reinsurance undertakings should ensure that the actuarial function knows and considers in its analysis the reliability of the sources of information and the consistency and stability of its process of collecting and publishing information over time."

This guideline is key as an important heterogeneity of corporate spread data among data providers is observed. It is strongly driven by the quality of the underlying methodologies as well as the completeness and the relevance of the bonds used to build benchmark spread indices.

The magnitude or the volatility of spreads can have adverse effects on the insurance balance sheet. This is particularly significant due to the influence of cost associated with options and guarantees on the liability side, coupled with bond reinvestment strategies on the asset side. To evaluate the impact of data quality, we have conducted a comparative analysis between a dataset sourced from S&P Global Market Intelligence[4] and an alternative dataset obtained from a different source. This alternative dataset is comparable to aggregated benchmark spread information commonly available in the financial market (called "benchmark data" in the rest of this paper).

This paper discusses the following topics:

1. Presentation of the typical risk-neutral credit models considered by insurers.
2. Construction of S&P Global Market Intelligence spread market data.
3. Comparison of the calibration of credit models based on two different sources of data.
4. Impact study of the credit data on the insurance liability valuation.

We demonstrate that the advanced methodology underpinning the S&P dataset allows it to outperform the benchmark data results by enhancing the performance and robustness of the credit models calibration and providing a more precise evaluation of scenario volatility. This, in turn, enables a more accurate assessment of solvency indicators.

Credits models

The Jarrow-Lando-Turnbull (JLT) and the Longstaff-Mithal-Neis (LMN) models are commonly used by insurance companies to model credit spread and default risk under a risk-neutral framework.

[1] Executive Director – Data, Valuations and Analytics – S&P Global Market Intelligence.
[2] Product Analysis and Design Director – Fixed Income Pricing – S&P Global Market Intelligence.
[3] EIOPA Guidelines on the valuation of technical provisions.
[4] Formerly IHS Markit.

The JLT model allows companies to model the credit rating transition through stochastic transition matrices while the LMN model focusses more on projecting stochastically the default intensity (also called hazard rate) of risky issuers. These two models are available within the Milliman ESG™.[5]

OVERVIEW OF THE JLT MODEL

The JLT model (see Jarrow, Lando, & Turnbull [1997]) is based on a historical annual transition matrix that is adjusted through a stochastic process π called a "risk premium adjustment," which captures the dynamics of the transition probabilities among different credit rating groups. π typically follows a Cox-Ingersoll-Ross (CIR) process under the risk-neutral probability measure:

$$\begin{cases} d\pi(t) = \alpha(\mu - \pi(t))dt + \sqrt{\pi(t)}\sigma dW_t \\ \pi(0) = \pi_0 \end{cases}$$

where W is a standard Brownian motion.

The JLT model relies on four parameters independent from the underlying ratings, namely, α the mean-reverting speed, μ the long-term level, σ the volatility parameter and π_0 the risk premium starting point.

OVERVIEW OF THE LMN MODEL

The LMN model (see Longstaff, Mithal & Neis [2005]) specifies the dynamics of the default intensity λ of each rating group G, as a CIR process under the risk-neutral measure:

$$d\lambda_G(t) = k_G\big(\theta_G - \lambda_G(t)\big)dt + \sigma_G\sqrt{\lambda_G(t)}dW_t^G$$

where $(W^G)_{G\in\mathcal{R}}$ are correlated Brownian motions, k_G is the mean-reverting speed, θ_G is the long-term level, σ_G is the volatility parameter and $\lambda_G(0)$ is the default intensity starting point.

Unlike the JLT, the LMN model embeds four parameters for each rating group.

For each group of issuers, logarithmic spreads can be directly priced using a closed form formula relying on four parameters, k_G, θ_G, σ_G and $\lambda_G(0)$. The parameters are estimated sequentially by decreasing rating quality in order to minimise the squared errors between the spreads induced by the model and the market spreads, consistently with the JLT target function.

CALIBRATION OF THE CREDIT MODELS

The typical input data considered for the calibration of credit spread and default models are market spreads relative to different maturities and different credit ratings. The JLT and LMN models both provide closed form formulae for the pricing of spreads in terms of their parameters.

As a result, the parameters can be estimated through a numerical optimisation procedure aiming at minimising the squared errors between market and model spreads.

In addition to the market spreads, the credit models require a loss given default (LGD) assumption. In the rest of this paper, for illustration purpose, the loss given default assumption has been set at 75%. However, in practice, the exact parameter may depend on each individual bond characteristic, with a particular emphasis on its seniority.

Besides, the JLT model takes as an input a historical transition matrix, usually provided by a credit rating agency. This matrix may be further adjusted in order to improve the accuracy of the risk-neutral calibration by increasing the default probabilities of the highest ratings compared to their historical values. Moreover, variants of the JLT model can be found in the literature, embedding additional parameters that aim at enhancing the final results.

Market spread data used for calibration

In this paper we address the importance of the quality of the market spread input data for the calibration of credit risk-neutral models. We compare two sets of data, namely data extracted from S&P Global Market Intelligence and alternative benchmark data obtained from a different source. In particular, S&P Global Market Intelligence provides enhanced data based on a multivariate factor model and tension-spline curve-fitting methodology used to calculate bond sector curves across attributes such as credit rating, country of risk, currency, sector and seniority.

ENSURING QUALITY OF INPUT DATA

Bond data used as an input to the multivariate factor model includes prices from S&P Global Market Intelligence's Corporate and Sovereign Bond Pricing Data service. A series of stringent data filtering and data cleaning is performed to ensure that outliers not representative of the market are not used in the calculations. In particular, the data-cleaning approach consists of defining bond buckets with similar characteristics and then eliminating bonds associated with yields too far from the bucket average yield.

FITTING OF A MULTIVARIATE FACTOR MODEL

Calculation of bond sector curves assumes that the bond market is governed by a fundamental factor model and decomposes observable bond prices into fundamental contributing factors. For example, suppose we know the yield for a German bond in the Consumer Services industry rated B. The factor model decomposes this yield into factor value contributions from the risks Germany country, B rating, and Consumer Services sector. In the model, the factor values are calibrated using a cross-sectional multivariate regression.

[5] More information is available at https://milliman.com/en/products/economic-scenario-generator.

Furthermore tension-spline functional forms are introduced into the multivariate regression to calibrate full-term structure of the factor values. For a given yield tenor T, the factor model can be written as:

$$yield(T) = \sum_i \beta_i f_i(T) + \varepsilon$$

The sum runs over all factors defined as part of the model and belonging to one of the following attributes: rating, currency, country of risk, sector, or tier (e.g., Financials, Consumer Services, BBB-, CCC+, SNRFOR, EUR, Germany etc.). Moreover:

- β_i denotes the i^{th} factor loading, which takes values 0 or 1 depending on whether the yield belongs to a bond that has the i^{th} factor. For example, for a German bond the US factor loading is 0.
- $f_i(T)$ is the i^{th} tenor dependent factor value function that has a tension-spline form allowing us to extend the regular factor model to a full-term structure factor model.
- ε is the error term of the factor model.

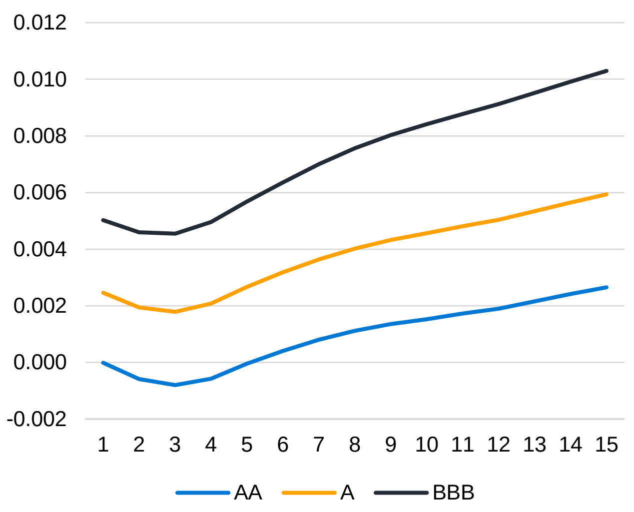
Quadratic programming is applied to determine the tension-spline coefficients along with proper boundary conditions by minimising the fitting error.

OBTAINING SPREAD CURVES DATA

Once the multivariate factor model has been calibrated, yield curves are obtained by summing the tenor dependent factor value functions relative to the bond characteristics. Spread curves are then deduced as the difference between the bond sector yield curves and the market swap curve (post-bootstrapping).

In this paper, we focus on euro corporate spreads for ratings AAA, AA, A, BBB, BB, B and CCC. The associated S&P Global Market Intelligence spread curves at 31 December 2021 are given below for AA, A and BBB ratings.

FIGURE 1: S&P GLOBAL MARKET INTELLIGENCE CORPORATE SPREADS



Finally, S&P Global Market Intelligence spread curves exhibit desirable features such as:

- The ordering of spreads according to the rating, in particular for low-quality ratings
- Smoothed term-structure in the tenor.

The following sections compare S&P Global Market Intelligence data to alternative benchmark data obtained from a different source that does not have the aforementioned properties.

Comparison of calibrations

This section sets out the calibration results as at 31 December 2021 obtained with S&P Global Market Intelligence data on one end, and with benchmark spreads on the other end. The results presented in this section have been produced using the Milliman ESG; they are illustrative and are not intended to be used for the valuation of an insurance company.

After the calibration of the credit models, we obtain several sets of parameters:

- $\widehat{\Theta} = (\alpha, \mu, \sigma, \pi_0)$ for the JLT model
- $\widehat{\Theta}_G := \left(k_G, \theta_G, \sigma_G, \lambda_G(0)\right)$, $G \in \{AAA, AA, A, BBB, BB, B, CCC\}$ for the LMN model.

There are therefore four (resp. 28) parameters to be estimated for the JLT (resp. LMN) model.

RESULTS FOR THE JLT MODEL

The table in Figure 2 presents the calibrated parameters for the JLT model, in two different configurations.

FIGURE 2: PARAMETERS FOR THE JLT MODEL

	π_0	α	μ	σ
Case A	67.00%	28.32%	126.72%	43.30%
Case B	98.84%	7.18%	222.33%	56.50%

Case A consists of calibrating the JLT model based on S&P Global Market Intelligence data.

Case B consists of calibrating the JLT model based on benchmark data.

In both cases, the calibration is based on the spreads associated with the AA, A and BBB ratings, which are those with the most liquid underlying data and the most represented in the portfolio of our French illustrative insurance company considered for the cash flow model impact study (see the impact study). This choice is also aligned with the typical mix observed in the French insurance industry, as the benchmark French national portfolio published by EIOPA for the computation of Volatility Adjustment yields at year-end 2021:

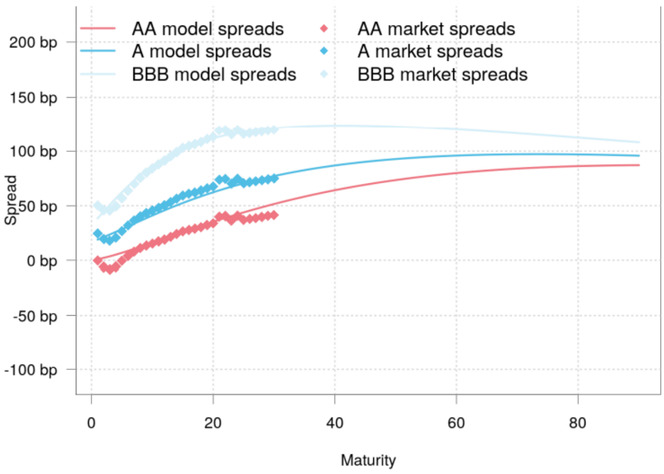
FIGURE 3: COMPOSITION CORPOPORATE PORTFOLIO (YEAR-END 2021)

	AAA	AA	A	BBB	BB	B	CCC
Illust. insurer	8%	20%	39%	31%	2%	0%	0%
EIOPA	12%	19%	41%	27%	1%	0%	0%

We note that the estimated parameters for Case A show lower long-term and initial levels as well as a higher mean-reverting speed. Furthermore, the volatility parameter is slightly lower when using S&P Global Market Intelligence data (Case A).

As an illustration, Figure 4 plots the theoretical and market spreads in Case A.

FIGURE 4: MODEL VS. MARKET SPREADS CURVES, CASE A



The model spreads are ordered by rating and the replication quality is acceptable.

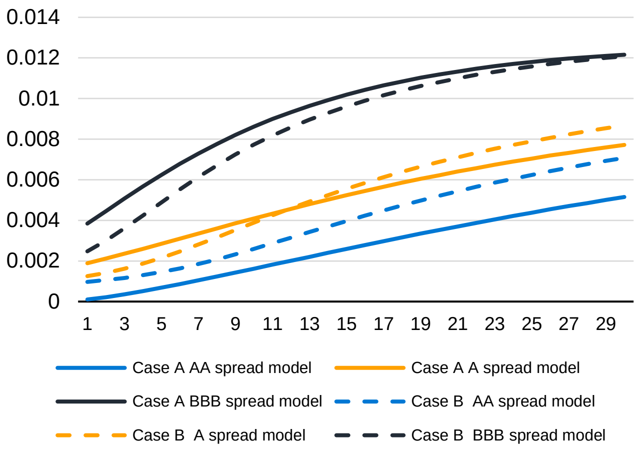
The fit quality of Case B is not as good as with the S&P Global Market Intelligence data. The table in Figure 5 shows the average absolute errors for the different Cases A and B; the aggregation along ratings is performed using the illustrative portfolio composition:

FIGURE 5: ABSOLUTE ERRORS FOR THE DIFFERENT SCENARIOS

	CASE A	CASE B
Average error	4.55 bp	8.94 bp

The graph in Figure 6 compares the model spreads calculated in both cases. We note that S&P Global Market Intelligence data calibration induces more differences in the ratings, as AA (resp. BBB) spread levels of Case A are below (resp. above) the ones of Case B.

FIGURE 6: CASE A VS. CASE B THEORICAL SPREADS CURVES



RESULTS FOR THE LMN MODEL

In the case of the LMN model, four parameters by rating have to be calibrated. We study hereafter the results associated with two distinct configurations:

- Case C consists of calibrating the LMN model on S&P Global Market Intelligence data.
- Case D consists of calibrating the LMN model on benchmark data.

The table in Figure 7 presents the associated calibrated volatility parameter σ_G for the LMN dynamics of ratings AA, A and BBB.

FIGURE 7: VOLATILITY PARAMETERS FOR THE LMN MODEL

	σ_{AA}	σ_A	σ_{BBB}
Case C	5.8%	5.2%	6.6%
Case D	7.2%	0.0%	6.7%

Furthermore, the table in Figure 8 shows the average absolute errors for the Cases C and D weighted with the illustrative insurer portfolio composition.

FIGURE 8: AVERAGE ABSOLUTE ERRORS FOR CASE C AND D

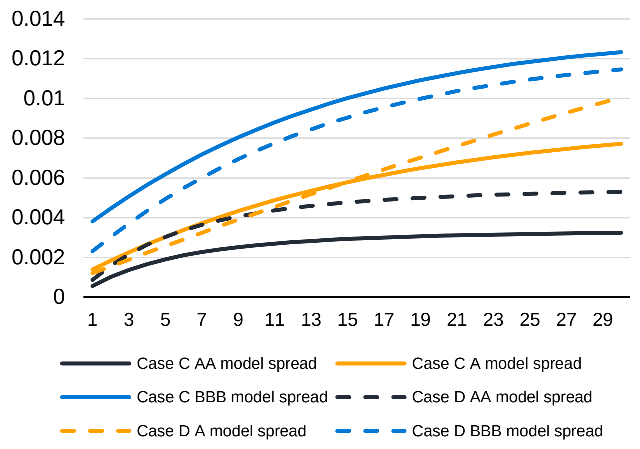
	CASE C	CASE D
Error	4.49 bp	4.62 bp

Compared to the JLT model, the improvement of the fit quality using S&P Global Market Intelligence data (Case C) is less pronounced. The use of the LMN model improved significantly the replication of the benchmark data. Nevertheless, this improvement comes at a cost; we will illustrate in the next section that the A spreads simulated in Case D are almost deterministic because the associated volatility parameter is 0%. This phenomenon is common with the LMN model when significant irregularities are observed in the data, as in Case D, inducing noise in the parameter's estimation.

S&P Global Market Intelligence data are smoother, providing a more “natural" shape for the model to be adjusted and improving both the efficiency and the robustness of the calibration.

Plotting the calculated model spreads in Cases C and D reveals an additional discrepancy in the data. Specifically, we observe a lack of ordering in the AA and A spreads when calibrated using the benchmark data:

FIGURE 9: SPREAD MODEL, CASE C VS. CASE D



Practitioners sometimes try to compensate for the abovementioned data issues by adding additional constraints to the LMN calibration, such as imposing that the mean-reverting speed and volatility parameters are ordered. Nevertheless, such approaches automatically result in a deterioration of the calibration errors. We note that S&P Global Market Intelligence data allows us to efficiently overcome these inconsistencies.

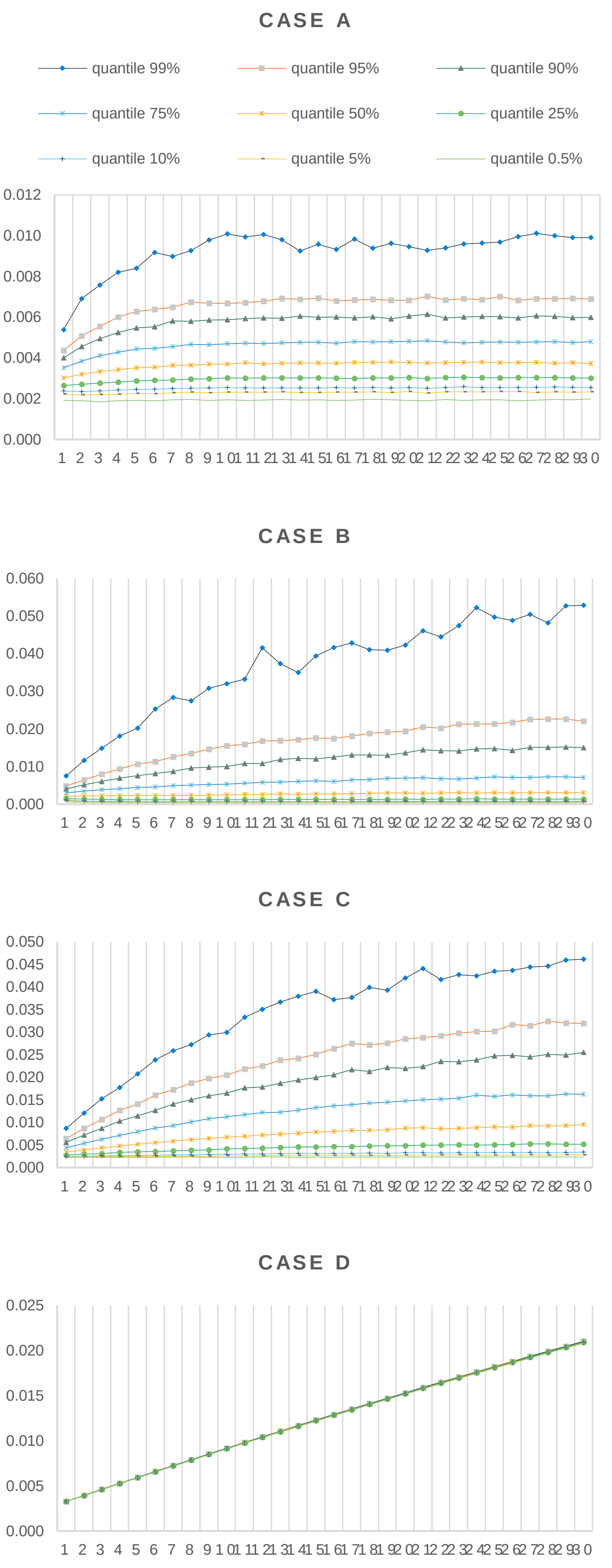
We have illustrated that the choice of the data provider directly impacts the performance and the robustness of the calibration process. In the following section we focus on the models simulation.

Comparison of simulations

The plots in Figure 9 show the spread diffusion cones in the different scenarios of Cases A, B, C and D. They are calculated from 3,000 simulations.

We chose as reference the spread rated A with a 5-year maturity as the most common rating within the portfolio we are analysing.

FIGURE 10: A 5-YEAR SPREAD DIFFUSION CONES



Regarding the JLT model, S&P Global Market Intelligence data implies less extensive scattering cones, explained by smaller volatility and long-term average parameters. In addition, we observe that the long-term spread level is reached sooner in Case A; this is due to the higher calibrated reversion speed.

In the case of the LMN model, the benchmark data lead to very thin diffusion cones, meaning that spreads are almost deterministic. As previously evoked, this observation is due to the data irregularities conducting to a null volatility parameter of the A default intensity process. Such a behaviour is problematic for the assessment of the time value of options and guarantees (TVOG) because the existence of optionality directly arises from the stochasticity of the underlying quantities.

Comparison of impacts

In this section, we present an impact analysis of the credit data sources and models (Cases A to D) in terms of Solvency II ratio for a typical French savings portfolio. The risk-neutral economic scenarios are generated by the Milliman ESG. These scenarios are then inputted in the cash flow model of an illustrative insurance company based on the typical features of a life insurance company in the French market. In particular:

1. The model refers to the Standard Formula (no internal model effects, no transitional measures etc.).
2. The actuarial valuation methodology used is “standard.” It might not reflect all the specificities of the various insurance companies in the market.
3. The assumptions are mostly derived from market data but are also based on our own knowledge of the French market.

Let us bear in mind that the primary objective of this paper is to examine the influence of credit market data rather than the impact of credit risk itself. The following example is presented purely for illustration purpose. In particular, it does not account for sovereign credit risk.

The table in Figure 11 details the key indicators, namely the Best Estimate of Liabilities (BEL), the Value of In-Force (VIF), the Solvency Capital Requirement (SCR) and the solvency ratio. The table shows the variations compared to the scenario where credit is not being modelled.

FIGURE 11: QUANTITIES OF INTEREST

	BE	VIF	SCR	SOLVENCY RATIO
Case A (JLT)	0.19%	-8.64%	-0.62%	-2.04%
Case B (JLT)	0.21%	-9.99%	-1.06%	-1.62%
Case C (LMN)	0.20%	-9.39%	-0.73%	-2.15%
Case D (LMN)	0.12%	-5.48%	-1.60%	1.36%

To conduct analysis, let us consider the following simplified breakdown of the insurance asset:

$$A = VIF + BE$$

As expected, in all cases, we observe that adding credit risk conducts to a decrease in the VIF and an increase in the BEL. Let us now focus on the decomposition of the BEL:

$$BEL = GB + FDB$$

In this context, the guaranteed benefits (GB) remain almost constant among all sensitivities while the contribution to the future discretionary benefits (FDB) increases. Indeed, GB are almost not impacted by the volatility of the economic environment and as such by the introduction of the credit risk. On the other end the FDB increases, which is a natural consequence of the decrease of the VIF. The increase of the FDB in the central scenario leads to a better loss absorption capacity of the SCR and then a lower SCR (net of loss absorption capacity).

We notice that Cases B and D (benchmark data) lead to a higher decrease in SCR than Cases A and C (S&P Global Market Intelligence data). This suggests that the use of the benchmark data may overestimate the decrease in the SCR. In turn, this implies an improvement in the solvency ratio for Cases B and D, in comparison to calibrations using S&P Global Market Intelligence data. This is particularly noteworthy and counterintuitive in Case D, because adding credit risk leads to a better solvency ratio.

Conclusion

This paper highlights the importance of the quality of spread market data for the calibration of credit risk-neutral models. In particular, the combination of both S&P Global Market Intelligence bond sector curves and the Milliman ESG has demonstrated the ability to deliver more comprehensive and reliable outcomes. Eventually, it should be acknowledged that considering high-quality data reduces the need for excessive constraints in the calibration process (e.g., ordering spread volatility in ratings).

REFERENCES

1. Jarrow, R. A., Lando, D., & Turnbull, S. M. (1997). A Markov model for the term structure of credit risk spreads. The Review of Financial Studies, 10(2), 481-523.
2. Longstaff, F. A., Mithal, S., & Neis, E. (2005). Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market. The Journal of Finance, 60(5), 2213-2253.

CONTACT

Pierre-Edouard Arrouy
pierre-edouard.arrouy@milliman.com

David Baranes
david.baranes@spglobal.com

Rashi Garg
rashi.garg@spglobal.com

For more information, please visit
milliman.com/en/products/economic-scenario-generator